

(19)



Europäisches Patentamt
European Patent Office
Office européen des brevets



(11)

EP 0 577 491 B1

(12)

FASCICULE DE BREVET EUROPEEN

(45) Date de publication et mention
de la délivrance du brevet:
17.03.1999 Bulletin 1999/11

(51) Int Cl.⁶: **G08B 13/196**

(21) Numéro de dépôt: **93401661.9**

(22) Date de dépôt: **28.06.1993**

(54) **Procédé et dispositif de surveillance d'une scène tridimensionnelle, mettant en oeuvre des capteurs d'imagerie**

Verfahren und Vorrichtung zur Überwachung einer dreidimensionalen Szene unter Verwendung von Bildsensoren

Process and device for monitoring a three-dimensional scene using imagery sensors

(84) Etats contractants désignés:
DE GB IT

(30) Priorité: **29.06.1992 FR 9207977**

(43) Date de publication de la demande:
05.01.1994 Bulletin 1994/01

(73) Titulaire: **MATRA CAP SYSTEMES**
78140 Velizy-Villacoublay (FR)

(72) Inventeur: **Bretaudeau, Francis**
F-76000 Rouen (FR)

(74) Mandataire: **Fort, Jacques**
CABINET PLASSERAUD
84, rue d'Amsterdam
75440 Paris Cedex 09 (FR)

(56) Documents cités:
EP-A- 0 402 829 **GB-A- 2 150 724**
GB-A- 2 194 367

- **ELECTRONICS AND COMMUNICATION ENGINEERING JOURNAL** vol. 2, no. 4, Août 1990, LONDON GB pages 155 - 164, XP223920
R.J. BLISSETT 'retrieving 3d information from video for robot control and surveillance'

Il est rappelé que: Dans un délai de neuf mois à compter de la date de publication de la mention de la délivrance du brevet européen, toute personne peut faire opposition au brevet européen délivré, auprès de l'Office européen des brevets. L'opposition doit être formée par écrit et motivée. Elle n'est réputée formée qu'après paiement de la taxe d'opposition. (Art. 99(1) Convention sur le brevet européen).

EP 0 577 491 B1

Description

[0001] La présente invention concerne la surveillance d'une scène éclairée, tridimensionnelle, ayant une structure géométrique connue ou déterminable, permettant notamment de détecter des anomalies de taille supérieure à une valeur déterminée (représentant par exemple des intrusions) ou, dans un mode évolué de réalisation, de localiser un mobile se déplaçant à l'intérieur de la scène.

[0002] On connaît déjà de nombreux procédés de surveillance (voir GB-A2 150 724 et R.J. Blissett : "Retrieving 3D information from video for robot control and surveillance", in Electronics and Communication Engineering Journal, pages 155-164, Août 1990, Vol.2 n°4) d'une scène éclairée à trois dimensions, l'éclairage étant dû à une source naturelle, telle que le soleil, ou à une source artificielle de rayonnement électro-magnétique ou d'ultrasons. Beaucoup de ces procédés font une comparaison entre l'image de la scène à un instant donné et l'image formée précédemment. Ces procédés ont notamment l'inconvénient de mal discriminer les modifications locales dues à une intrusion de celles dues à une modification locale d'éclairage. De façon générale, ces procédés et dispositifs ont pour la plupart une sélectivité insuffisante pour permettre de différencier, avec une probabilité élevée, des anomalies caractérisées par leur taille, et notamment leur saillie par rapport à la scène d'origine, donc de différencier par exemple un rongeur et un être humain.

[0003] La présente invention vise à fournir un procédé et un dispositif de surveillance ou de contrôle d'une scène, répondant mieux que ceux antérieurement connus aux exigences de la pratique, notamment en ce qu'ils assurent une sélectivité élevée en ne mettant en oeuvre que des moyens matériels simples.

[0004] Dans ce but l'invention propose notamment un procédé de surveillance d'une scène éclairée à trois dimensions, suivant lequel :

- on constitue un modèle géométrique tridimensionnel (dit 3D) de référence de ladite scène associant, à chaque point de la scène, des points homologues dans le champ d'au moins deux capteurs placés à distance l'un de l'autre et sensibles à un paramètre radiométrique (tel que la brillance) de la scène ;
- on synthétise périodiquement la scène réelle sur le modèle géométrique tridimensionnel, à partir de chacune des représentations qui en sont fournies par les capteurs ; et, à partir de là,
- on compare les scènes synthétisées entre elles pour faire apparaître les écarts dus à une modification du modèle 3D qui ont une taille supérieure à une valeur déterminée.

[0005] On voit que le modèle de référence est purement géométrique et n'intervient que pour permettre d'établir la correspondance entre points homologues dans les champs des différents capteurs. La comparaison peut s'effectuer entre des images prises à des instants identiques, donc dans des conditions d'éclairage qui sont les mêmes. Etant donné que la comparaison est effectuée entre des scènes synthétisées, une modification des conditions extérieures modifiant les valeurs radiométriques (brillance ou couleur) d'un point sera sans effet sur la comparaison, à condition que les capteurs aient des caractéristiques de réponse similaires.

[0006] Il faut incidemment noter que le terme "scène" ne doit pas être interprété comme limité au cas d'une zone continue, correspondant à la zone de recouvrement des champs des capteurs. La scène peut être constituée de portions discrètes dans les champs, seules jugées importantes et à surveiller.

[0007] La comparaison des scènes synthétisées peut être effectuée par détermination des coefficients d'intercorrélation entre les deux scènes synthétisées, sur des voisinages déterminés des points de la scène, et par comparaison de ces coefficients à un seuil qui sera évidemment toujours inférieur à l'unité.

[0008] Le modèle géométrique 3D de la scène peut être constitué à partir des images de la scène, dans un état de référence, fournies par les capteurs, mémorisées soit pixel par pixel soit, plus généralement, à l'aide d'éléments géométriques tels que des facettes.

[0009] Le modèle géométrique peut également être établi directement sous forme mathématique à partir d'une connaissance du terrain fournie par des observations antérieures ou à l'aide de cartes et les modèles de représentation géométrique en deux dimensions par les capteurs peuvent en être déduites à partir de la connaissance de la position des capteurs par rapport à la scène.

[0010] Dans la plupart des cas, on sera amené à donner aux capteurs un écartement mutuel inférieur à la distance entre les capteurs et les points de la scène qui en sont les plus proches : le rapport entre l'écartement des capteurs et la distance sera un paramètre prépondérant pour sélectionner la saillie minimale que doit avoir un objet par rapport à la scène d'origine pour être détecté.

[0011] Les opérations de détection définies ci-dessus peuvent être répétées, par exemple à la cadence de prise de vue des capteurs, qui seront généralement des caméras. Ainsi on peut suivre les évolutions d'un objet dans la scène.

[0012] L'invention propose également un dispositif de surveillance permettant de mettre en oeuvre le procédé ci-dessus défini. Ce dispositif comprend plusieurs capteurs sensibles à un paramètre radiométrique, dans des longueurs d'onde comparables, comprenant des moyens de mémorisation d'un modèle purement géométrique tridimensionnel

de référence d'une scène à surveiller, associant, à chaque point de la scène, des points homologues dans le champ des capteurs ; des moyens permettant, à partir du modèle géométrique et des signaux représentatifs de la valeur radiométrique de chaque pixel d'image fourni par les capteurs, de synthétiser la scène réelle par projection sur le modèle géométrique tridimensionnel ; et des moyens de comparaison de voisinages de dimension déterminée des scènes synthétisées, permettant de déterminer les écarts entre eux.

[0013] L'invention est susceptible de nombreux modes de mise en oeuvre différents, permettant de l'adapter aussi bien au cas où l'on peut disposer ou obtenir un modèle géométrique tridimensionnel sous forme d'un modèle numérique de terrain qu'au cas où il est nécessaire de déterminer une transformation établissant, au moins de façon approchée, une correspondance entre chaque point du champ d'un des capteurs avec le point homologue du champ de l'autre capteur.

[0014] L'invention sera mieux comprise à la lecture de la description qui suit de modes particuliers de réalisation, donnés à titre d'exemple non limitatifs. La description se réfère aux dessins qui l'accompagnent, dans lesquels :

- la figure 1 montre schématiquement la disposition de deux capteurs par rapport à une scène et l'effet d'une anomalie en relief sur la scène, dans le cas d'un dispositif de surveillance que l'on peut qualifier de "fond de terrain" ;
- la figure 2 est un schéma destiné à faire apparaître l'influence de différents paramètres sur la détection ;
- la figure 3 est un organigramme d'un mode de réalisation du procédé ;
- la figure 4 est un schéma de principe montrant une constitution possible d'un dispositif ;
- les figures 5A et 5B montrent des dispositions particulières, avantageuses du point de vue du traitement, des images fournies par les caméras constituant les capteurs d'un dispositif ;
- la figure 6 montre une architecture possible de la chaîne de traitement d'un dispositif.

[0015] Avant de décrire de façon détaillée les étapes d'un procédé suivant l'invention de détection d'intrusions, on définira l'approche générale mise en oeuvre conformément à l'invention, dans le cas particulier d'une scène constituée par un terrain dont il est possible de représenter la structure par un modèle géométrique sous forme d'un modèle numérique de terrain (en abrégé MNT), donnant les coordonnées de chaque point du terrain dans un système d'axes fixe x , y et z .

[0016] La figure 1 montre un dispositif n'ayant que deux capteurs 10 et 12, constitués par exemple par des caméras sensibles dans le domaine visible, qui ont des champs angulaires qui se recouvrent dans la scène à observer. Dans l'image fournie par un capteur d'ordre i (i étant égal à 1 ou 2 dans le cas considéré) un point du modèle numérique de terrain ayant des coordonnées (x, y, z) dans le système d'axe référentiel choisi correspondra à un point de coordonnées (α_i, β_i) et de radiométrie S_i dans un référentiel à deux dimensions qui est associé cette fois au capteur respectif. Il est possible de déterminer, par le calcul ou des mesures préalables, les transformations permettant de passer des coordonnées et de la radiométrie dans un référentiel lié à la scène aux coordonnées et à la radiométrie dans un référentiel lié à chaque capteur.

[0017] Connaissant le modèle géométrique 3D de la scène à observer et la répartition des valeurs radiométriques données par un des capteurs, on peut construire une scène de synthèse associée à ce capteur, en réaffectant à chaque point connu dans la scène, c'est-à-dire à chaque point de coordonnées (x, y, z) , la valeur radiométrique fournie par le capteur.

[0018] On constitue ainsi, pour chaque capteur, une scène de synthèse en associant la géométrie réelle, fournie par le MNT, et la valeur radiométrique, soit directement fournie par le capteur, soit soumise à un traitement préalable, par exemple de filtrage.

[0019] Si la géométrie 3D de la scène réelle n'a pas subi de modification dans une zone déterminée, les scènes synthétisées resteront quasi-identiques l'une à l'autre pour cette zone, même s'il y a eu évolution de la radiométrie réelle. Si même, du fait par exemple d'une réponse différente des deux capteurs et des variations d'éclairement, les scènes synthétisées ne restent pas strictement identiques, les scènes synthétisées resteront dans tous les cas étroitement corrélées, en l'absence de modification de la géométrie du terrain.

[0020] Si en revanche la scène observée a subi des modifications géométriques par rapport au modèle connu, par suite d'une intrusion, la valeur radiométrique correspondant à un point de la perturbation, vu d'un des capteurs, sera projetée lors de la synthèse sur une position géométrique du point d'où elle n'est pas issue, étant donné que le dispositif ne connaît que le modèle géométrique n'ayant pas cette perturbation. Il en sera de même pour la valeur radiométrique fournie, pour le même point de la perturbation, par l'autre capteur. En conséquence les scènes de synthèse élaborées par deux capteurs, chacun à partir des informations radiométriques qu'il reçoit et du modèle géométrique de référence, divergeront. Les écarts entre les deux scènes de synthèse seront fonction de la hauteur et du volume de la perturbation.

[0021] Cette influence de la hauteur de la perturbation apparaît immédiatement si l'on considère le cas particulier d'une perturbation n'ayant qu'une très faible épaisseur : dans ce cas il y a uniquement modification des valeurs radiométriques dans une zone particulière de la scène réelle. Cette scène restera géométriquement identique au modèle de référence. La perturbation, comme une perturbation qui serait due à une modification locale d'éclairement (passage

d'une ombre portée par exemple) ne sera pas détectée comme une intrusion.

[0022] Le dispositif peut en conséquence être robuste aux variations de conditions d'éclairage et être insensible au passage d'objets de petite taille sur le sol.

[0023] On verra plus loin que la taille, et notamment la hauteur, minimale détectée dépend notamment de la résolution des capteurs et de la distance entre eux.

[0024] On voit également sur la figure 1 que l'anomalie 14 montrée en tirets se traduit par une modification de l'image fournie par le capteur 10 sur la largeur l , ce qui peut se traduire comme l'adjonction d'une ombre sur l'image. La même anomalie se traduit pour l'image fournie par le capteur 12, par une modification sur la largeur l' : les scènes de synthèse ne coïncideront donc plus sur l'ensemble des "ombres". Cela reste vrai quelle que soit la nature de l'éclairage, qu'il soit optique, électro-magnétique à fréquence élevée ou même ultrasonore.

[0025] La sensibilité de détection du dispositif dépend de la résolution au niveau où sont susceptibles de se produire des intrusions. Si on désigne par exemple par \underline{D} l'intervalle entre deux capteurs, supposés identiques, et par $E_i = P_i$ la résolution au niveau du terrain, pour le point i et le capteur 12, une anomalie sera détectée à la seule condition qu'elle déborde des domaines de résolution angulaire des deux capteurs, associés à un même emplacement du modèle de terrain. En conséquence, l'anomalie indiquée en 14₁ ne sera pas détectée, alors que les anomalies 14₂ et 14₃ peuvent l'être. En conséquence, on peut déterminer les conditions à remplir pour détecter une anomalie dès qu'elle dépasse une hauteur déterminée au-dessus du sol.

[0026] Si en particulier le terrain est loin des capteurs et n'est pas trop perturbé, on peut considérer que la hauteur minimale \underline{d} par rapport au terrain que doit présenter un obstacle pour être détecté est reliée aux autres paramètres par :

$$D/(L-d) = (L \cdot \text{pix})/d$$

où L est la distance moyenne des capteurs au terrain, D est la distance entre les capteurs, et pix désigne la résolution angulaire correspondant à la résolution au sol E_i .

[0027] Dans le cas d'un dispositif que l'on peut qualifier de "fond de terre" destiné à détecter des intrus au sol, tout en évitant les fausses alarmes dues à des animaux de petite taille, L sera toujours très supérieur à \underline{d} . On peut alors écrire :

$$L^2 = D \cdot d / \text{pix}$$

[0028] A titre d'exemple, on peut adopter, pour détecter des objets de taille supérieure à 10 centimètres, $d=20$ cm.

[0029] Le résultat peut alors être atteint avec :

$$\text{pix} = 10^{-3} \text{ rad}$$

$$L = 20 \text{ m}$$

$$D = 2 \text{ m}$$

[0030] Quelle que soit la solution adoptée, on voit qu'on utilise la connaissance de la structure géométrique de la scène pour donner à la détection un caractère de comparaison géométrique, ce qui augmente les performances. La connaissance que l'on a a priori élimine l'influence des éléments connus et favorise la détection des modifications géométriques dues aux intrusions.

[0031] Le dispositif peut être complété par des capteurs d'autres natures, ou mettre en oeuvre plusieurs systèmes associés travaillant dans des longueurs d'onde différentes. De plus il peut mettre en oeuvre, en plus d'une détection faisant appel à une comparaison entre valeurs radiométriques fournies par deux capteurs ou davantage (par exemple en infra-rouge), une détection temporelle utilisant une comparaison entre images successives.

[0032] Enfin, une fois un objet détecté, il est possible de le faire disparaître en incorporant ses caractéristiques dans le modèle géométrique de la zone observée, soit par une nouvelle prise de vue et une analyse des nouvelles vues, soit par voie purement mathématique.

[0033] Le procédé, tel qu'il vient d'être décrit jusqu'ici, implique la connaissance du modèle géométrique 3D de la scène observée pour élaborer les scènes de synthèse.

[0034] Mais il suffit, pour mettre en oeuvre le procédé, de disposer d'une correspondance géométrique entre les pixels des images des différents capteurs. Cette correspondance géométrique est établie pour la scène de référence afin de déceler ensuite les perturbations. Toute perturbation géométrique de la scène de référence provoque une différence radiométrique entre les points associés par la correspondance dans la scène de référence. Comme on le verra plus loin, cette correspondance peut être établie à partir de l'identification dans l'image de points d'amer dans les champs des différents capteurs. Enfin, un modèle peut être établi en assimilant le terrain à une mosaïque de facettes ayant chacune une forme géométrique simple.

[0035] Dans ces conditions, à partir de l'approche exposée ci-dessus, divers procédés de détection sont susceptibles

d'être mis en œuvre. On mentionnera maintenant deux solutions, qui diffèrent par la méthode de mise en correspondance géométrique du modèle et de l'image synthétisée.

[0036] La première méthode suppose qu'un modèle numérique du terrain a été réalisé et mémorisé. Ce modèle numérique peut être constitué à partir de photographies aériennes, de cartes, ou même de mesures sur le terrain. A partir de la connaissance de ce modèle mathématique, on peut déterminer par le calcul, pour une position donnée de chacun des capteurs, au nombre de deux au moins pour chaque scène, les transformations qui font correspondre, à chaque point de coordonnées (x,y,z) du modèle numérique, les coordonnées α_i , β_i de l'image de ce point dans le repère associé au capteur d'ordre i.

[0037] En d'autres termes, on peut écrire, en désignant par \mathbf{T} les transformées déterminables :

$$\alpha_i = \mathbf{T} \alpha^i(x,y,z)$$

$$\beta_i = \mathbf{T} \beta^i(x,y,z) \quad (1)$$

[0038] Le signal radiométrique $S(\alpha_i, \beta_i)$ fourni en chaque point de l'image du capteur i est proportionnel au flux réémis par l'élément de surface associé aux points (x,y,z) du modèle, le coefficient de proportionnalité étant sensiblement constant sur un voisinage de quelques pixels.

[0039] Le schéma de détection, dans le cas de deux capteurs 10 et 12 constitués par deux caméras, est alors celui montré en figure 3. A un point (x,y,z) diffusant un flux \underline{S} , on peut associer les points image (α_1, β_1) de la caméra 1 et (α_2, β_2) de la caméra 2. On peut également en déduire les signaux S1 et S2.

[0040] On peut donc ainsi créer deux objets de synthèse correspondant chacun à une des caméras, constitués par des points ayant chacun :

- des coordonnées (x,y,z) qui ne varient pas dans le temps, puisqu'elles correspondent à la géométrie réelle de la scène,
- des radiométries qui fluctuent suivant les conditions d'éclairement.

[0041] L'opération de détection s'effectue par comparaison des signaux S1 et S2 associés à un même point (x,y,z) et cela pour chacun des points (x,y,z).

[0042] Cette comparaison peut s'effectuer par diverses opérations mathématiques. Une solution donnant de bons résultats consiste à calculer le coefficient C(p) d'intercorrélation normalisé local entre les deux modèles de synthèse, sous la forme :

$$C(p) = \frac{\sum S_1(T_\alpha^1(P'), T_\beta^1(P')) S_2(T_\alpha^2(P'), T_\beta^2(P')) p'_e v(p)}{[\sum S_1(T_\alpha^1(P'), T_\beta^1(P'))^2 \times \sum S_2(T_\alpha^2(P'), T_\beta^2(P'))^2]^{1/2} p'_e v(p)}$$

où

V(p) désigne un voisinage des points P respectifs, en général sur 3x3 ou 5x5 pixels dans le cas d'un réseau carré ;
p' désigne un point courant de V(p), de coordonnées (x',y',z').

[0043] Le coefficient d'intercorrélation normalisé est alors égal à 1 si, pour un voisinage V(p) donné, on a :

$$S_1(T_\alpha^1(p'), T_\beta^1(p')) = \lambda S_2(T_\alpha^2(p'), T_\beta^2(p'))$$

[0044] Cette égalité est respectée si la scène observée n'est pas perturbée.

[0045] Cette méthode permet de s'affranchir des conséquences de différences de gains entre les capteurs et des effets dus à des conditions angulaires distinctes de prise de vue, ayant une influence sur la rétrodiffusion, vers les capteurs.

[0046] Dans la mesure où la scène observée est parfaitement diffusante, il suffit d'une comparaison pixel à pixel des deux images.

[0047] Dans le cas d'une intrusion, le modèle de synthèse est reconstruit avec une radiométrie qui n'est plus celle associée à la géométrie de la zone observée. On associe en effet, à un point de coordonnées (x,y,z) masqué par

l'intrus, une radiométrie qui n'est pas issue du point de la scène de référence. Et cela est vrai pour tous les points masqués par la perturbation.

[0048] La détection est déterminée en émettant un signal d'alarme lorsque le coefficient de corrélation est inférieur à un seuil s , qui est évidemment choisi toujours inférieur à 1.

[0049] Une telle technique de détection a l'avantage d'être peu sensible aux variations de dynamique et notamment d'éclairement, ce qui constitue un facteur de robustesse important et permet d'omettre tout étalonnage radiométrique.

[0050] Une seconde approche doit être utilisée lorsque l'on ne connaît pas initialement un modèle numérique de terrain : elle permet de revenir aux conditions initiales du cas où un MNT est disponible.

[0051] Une solution consiste à utiliser des points remarquables dans la scène, facilement repérables dans chaque image et dont on connaît les coordonnées (x, y, z) dans un même repère géographique, constituant des points d'amer. On désignera par $\{x_k, y_k, z_k\}$ les coordonnées du point d'amer d'ordre k et par :

- $\{\alpha_{k1}, \beta_{k1}\}$ les coordonnées des images du point d'ordre k dans le repère lié à la caméra 1,
- $\{\alpha_{k2}, \beta_{k2}\}$ les coordonnées des images dans le repère lié à la caméra 2.

[0052] On recherche alors la transformation permettant de passer des coordonnées dans la scène réelle aux coordonnées dans l'image :

$$\alpha_i = T_{\alpha i}(x, y, z)$$

$$\beta_i = T_{\beta i}(x, y, z).$$

[0053] Suivant la complexité de la scène, on sera amené à adopter des équations de formes plus ou moins complexes pour les transformations T . Le nombre de points d'amer nécessaire pour donner une représentation suffisamment exacte augmentera avec la complexité de la transformation. Lorsque la scène n'est pas trop tourmentée, il est possible d'adopter des transformations T_α et T_β de forme bilinéaire :

$$T_\alpha(x, y, z) = (\alpha_{000}) + (\alpha_{100})x + (\alpha_{010})y + (\alpha_{001})z + (\alpha_{110})xy + (\alpha_{101})xz + (\alpha_{011})yz$$

$$T_\beta(x, y, z) = (\beta_{000}) + (\beta_{100})x + (\beta_{010})y + (\beta_{001})z + (\beta_{110})xy + (\beta_{101})xz + (\beta_{011})yz$$

[0054] Le calcul des coefficients $\alpha_{000}, \dots, \beta_{011}$ s'effectue en écrivant, pour l'ensemble des points d'amer, les équations :

$$\alpha_k^i = T_{\alpha}^i(x_k, y_k, z_k)$$

$$\beta_k^i = T_{\beta}^i(x_k, y_k, z_k)$$

[0055] On obtient ainsi des systèmes d'équations linéaires qui peuvent être résolus sans difficulté : dans le cas ci-dessus sept coefficients sont à déterminer, ce qui exige au moins sept points d'amer. Un nombre plus important peut être utilisé, de façon à retenir des coefficients représentant un optimum.

[0056] Une fois les transformations T_α et T_β connues, on se retrouve dans le cas précédent, puisque l'on connaît l'équivalent d'une MNT.

[0057] Une autre solution consiste à utiliser une grille de transformation, projetée sur la scène réelle et dont on repère la représentation dans chaque image. Cette grille peut notamment être constituée par un corroyage, obtenu par exemple par balayage à l'aide d'un laser et repérage, à chaque impulsion laser, du point de l'image représentant l'impact sur le terrain.

[0058] On décrira maintenant, en faisant référence à la figure 4, la constitution de principe d'un dispositif de surveillance, destiné à la détection automatique d'intrusion dans des zones connues. La figure 4 montre deux capteurs, constitués par des caméras 10 et 12 dont les champs présentent une zone de recouvrement 16 dont une partie ou la totalité constitue la scène surveillée. L'éclairage peut être naturel ; de nuit ou pour améliorer la sensibilité de détection à courte distance, l'éclairage naturel peut être remplacé ou complété par une source artificielle, visible ou infra-rouge.

Dans le cas de détection de personne, il suffit d'une résolution au sol de 10x10 cm et d'une cadence de deux détections par secondes, qui correspond à un déplacement de 1,5 m pour une personne à 10 km/h. Les capteurs peuvent être des caméras infra-rouge à CCD ayant une matrice de 600x700 pixels. Les deux caméras 10 et 12 sont synchronisées, par exemple en commandant la caméra 12 à partir du signal de synchronisation de la caméra 10.

[0059] Les signaux vidéo provenant des caméras sont appliqués à une unité de traitement 18 qui peut être regardée comme comprenant :

- deux cartes d'acquisition 20 et 22 permettant de récupérer la partie utile des images des caméras,
- une carte processeur de signal 24,
- une carte de mémoire vive 26, et, éventuellement,
- une carte de communication 28 permettant de connecter l'unité de traitement à un réseau téléphonique.

[0060] Bien que l'on puisse disposer les caméras 10 et 12 dans n'importe quelle orientation relative, il est cependant avantageux, pour réduire la complexité du traitement et la capacité de mémoire requise, de respecter des contraintes qui dans la pratique ne sont pas gênantes et dont certaines seront maintenant exposées.

[0061] On rappellera tout d'abord que les pixels correspondants de deux images fournies d'une même scène par deux caméras sont sur deux droites, dites épipolaires, passant chacune par le centre optique O_1 et O_2 de l'image respective : dans chaque image, le pixel correspondant au même point P de la scène est dans l'image -constitué par une droite épipolaire- du plan $O_1 P O_2$. Les épipolaires ont la propriété que tout pixel de l'épipolaire dans une image a un pixel homologue dans l'épipolaire correspondante de l'autre image si les repères de référence des caméras restent inchangés.

[0062] Une configuration intéressante est de placer les caméras de façon que leurs repères (directions ligne et trame) soient parallèles et que les points O_1 et O_2 soient situés sur une même droite parallèle à la direction de ligne. Deux épipolaires telles que 30₁ et 30₂ sont alors des droites parallèles à la direction de ligne (figure 5A). Les traitements peuvent alors s'effectuer ligne à ligne suivant les épipolaires : il suffit de mémoriser dans la carte 26 les lignes correspondant aux épipolaires des zones observées et un nombre de lignes en-dessus et au-dessous correspondant à l'étendue du voisinage choisi pour le calcul de corrélation.

[0063] Lorsqu'il n'est pas possible ou peu commode d'adopter la disposition de la figure 5A, celle de la figure 5B est encore avantageuse. Les épipolaires 32₁ et 32₂ sont alors des droites inclinées, mais faisant le même angle par rapport aux lignes des deux images.

[0064] Il faut alors effectuer un traitement supplémentaire, en comparaison de la figure 5A. Ou bien on sélectionne les points homologues des épipolaires et des parallèles aux épipolaires appartenant au voisinage choisi, ou bien on augmente la taille de la mémoire 26.

[0065] La figure 6 montre une constitution possible de la partie traitement d'un dispositif dans le cas d'un modèle de terrain à l'aide d'une grille mémorisée. On supposera que les centres optiques des caméras 10 et 12 sont disposés au voisinage d'une même ligne droite parallèle au sens de balayage dans la direction ligne. Dans ce cas la mémoire 26 (figure 4) peut comporter deux espaces mémoire 26₁ et 26₂ correspondant à un paquet de quelques lignes de pixels seulement, centré sur la ligne d'ordre j dont les pixels sont voisins. Le processeur peut être regardé comme ayant deux voies de traitement 24₁ et 24₂ effectuant les projections et recevant, sur des entrées respectives 34₁ et 34₂, les modèles géométriques sous forme numérique.

[0066] Le processeur comporte également une unité 36 de comparaison des projections, généralement par détermination d'un facteur d'intercorrélation sur chaque voisinage et de comparaison à un seuil.

[0067] La sortie de l'unité 36 est appliquée à un bloc 38 de transfert des images vers un appareil d'enregistrement et éventuellement un appareil de visualisation 40. Le processeur 36 comporte également une sortie 40 commandant le passage d'un paquet de lignes centré sur la ligne j à un paquet centré sur la ligne i+1 une fois la comparaison terminée. En d'autres termes, le processeur pilote de façon asynchrone la mémorisation des paquets de lignes.

[0068] L'architecture montrée en figure 6 permet de limiter la quantité de mémoire vive à écriture et lecture rapide.

[0069] L'invention est susceptible de nombreuses variantes de réalisation et de nombreuses applications autres que celles qui ont été spécifiquement évoquées.

[0070] Par exemple le dispositif peut être utilisé pour localiser un robot à l'intérieur de la scène, le robot constituant une anomalie.

[0071] Dans ce cas, il peut être avantageux de placer les capteurs sur un organe de déplacement (chariot de pont roulant par exemple) muni de moyens de mesure permettant de disposer en permanence des coordonnées des capteurs par rapport au terrain. Il est ainsi possible, à partir d'une mémoire centrale stockant un modèle numérique de terrain correspondant à la totalité du domaine d'évolution du robot, de mettre en permanence à jour la mémoire du processeur de traitement pour qu'elle corresponde à l'emplacement des capteurs. Le dispositif suivant l'invention a l'avantage d'être exempt de la plupart des défauts des systèmes de visualisation montés sur les robots eux-mêmes, notamment la limitation du champ de vision en distance et l'impossibilité d'observer le domaine de déplacement du

robot derrière des obstacles.

Revendications

- 5 1. Procédé de surveillance d'une scène tridimensionnelle, ayant une structure géométrique connue ou déterminable, permettant de détecter l'apparition d'une anomalie en relief, incluant les étapes suivantes :
 - 10 - on constitue et on mémorise un modèle géométrique tridimensionnel de référence de ladite scène, dans un état de référence, associant, à chaque point de la scène, des points homologues dans le champ d'au moins deux capteurs (10, 12) placés à distance l'un de l'autre et sensibles à un paramètre radiométrique (tel que la brillance) de la scène ;
 - on synthétise périodiquement la scène réelle sur le modèle géométrique tridimensionnel, à partir de chacune des représentations qui en sont fournies par les capteurs, pour obtenir autant de scènes synthétisées que de
 - 15 - on compare des voisinages appartenant aux scènes synthétisées entre eux pour faire apparaître les écarts dus à une modification du modèle géométrique tridimensionnel, en ne retenant que les écarts correspondant à une anomalie de hauteur supérieure à une valeur déterminée minimale.
- 20 2. Procédé selon la revendication 1, caractérisé en ce qu'on compare les scènes synthétisées par détermination des coefficients d'intercorrélation entre les deux scènes synthétisées, sur des voisinages déterminés des points de la scène, et en ce qu'on compare les coefficients obtenus à un seuil.
- 25 3. Procédé selon la revendication 1 ou 2, caractérisé en ce qu'on constitue le modèle géométrique de la scène à partir des images de la scène, dans un état de référence, fournies par les capteurs, mémorisées pixel par pixel.
4. Procédé selon la revendication 1 ou 2, caractérisé en ce qu'on établit le modèle géométrique sous forme mathématique à partir d'une connaissance du terrain fournie par des observations antérieures ou à l'aide de cartes.
- 30 5. Procédé selon la revendication 1 ou 2, caractérisé en ce qu'on constitue le modèle géométrique de la scène à partir de points d'amer de coordonnées (x,y,z) connues et d'une équation de forme prédéterminée.
6. Procédé selon la revendication 1 ou 2, caractérisé en ce que l'on constitue le modèle géométrique à partir d'une grille de corroyage, réalisée sur la scène par balayage à l'aide d'un laser et enregistrée sous forme du repérage
- 35 par chaque capteur, à chaque impulsion laser, du point de l'image représentant l'impact sur le terrain.
7. Procédé selon l'une quelconque des revendications précédentes, caractérisé en ce qu'on compare de plus des images successives fournies par les capteurs pour réaliser une détection temporelle.
- 40 8. Procédé selon l'une quelconque des revendications précédentes, caractérisé en ce que les capteurs ont un écartement mutuel inférieur à la distance entre les capteurs et les points les plus proches de la scène.
9. Dispositif de surveillance d'une scène éclairée tridimensionnelle ayant une structure géométrique connue ou déterminable, permettant de détecter des anomalies en relief sur la scène, comprenant au moins deux capteurs
- 45 (10,12) sensibles à un paramètre radiométrique, dans des longueurs d'onde comparables, et comprenant également : des moyens de mémorisation d'un modèle purement géométrique tridimensionnel de référence de la scène à surveiller, associant, à chaque point de la scène, des points homologues dans le champ des capteurs ; des moyens permettant de synthétiser des scènes pour obtenir autant de scènes synthétisées que de capteurs, chacune des scènes étant synthétisée à partir du modèle géométrique et des signaux représentatifs de la valeur
- 50 radiométrique de chaque pixel d'image fourni par un des capteurs par projection sur le modèle géométrique tridimensionnel ; et des moyens de comparaison de voisinages de dimension déterminée sur l'ensemble des scènes synthétisées à partir des signaux fournis par les différents capteurs, permettant de déterminer les écarts entre les mêmes voisinages dans les deux scènes synthétisées.
- 55 10. Dispositif selon la revendication 9, caractérisé en ce que lesdits capteurs comportent de plus des capteurs travaillant dans des plages de longueur d'onde différentes des deux dits capteurs.
11. Dispositif selon la revendication 10 de surveillance destiné à localiser un robot mobile dans la scène, caractérisé

en ce que lesdits capteurs sont portés par des moyens de déplacement par rapport à la scène et sont associés à une mémoire à mise à jour du modèle géométrique mémorisé en fonction de la position des capteurs par rapport à la scène.

Patentansprüche

1. Verfahren zur Überwachung einer dreidimensionalen Szene mit einer bekannten oder bestimmbaren geometrischen Struktur, welches die Erfassung des Auftretens einer Reliefanomalie ermöglicht, mit den folgenden Schritten:
 - Erstellen und Speichern eines dreidimensionalen geometrischen Bezugsmodells der Szene in einem Bezugszustand, indem jedem Punkt der Szene homologe Punkte in dem Feld von mindestens zwei Sensoren (10, 12) zugeordnet werden, welche in einem Abstand voneinander angeordnet sind und in Bezug auf einen radiometrischen Parameter (wie etwa die Helligkeit) der Szene empfindlich sind;
 - periodisches Synthetisieren der realen Szene auf dem dreidimensionalen geometrischen Modell ausgehend von jeder der von den Sensoren gelieferten Darstellungen davon, zum Herstellen einer Anzahl von synthetisierten Szenen, welche gleich derjenigen der Sensoren ist; und
 - Vergleichen von den synthetisierten Szenen zugehörigen Umgebungen miteinander, um Abweichungen infolge einer Modifikation des dreidimensionalen geometrischen Modells erkennbar zu machen, wobei nur diejenigen Abweichungen aufbewahrt werden, welche einen vorgegebenen Minimalwert übersteigenden Höhenanomalie entsprechen.
2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß die synthetisierten Szenen durch die Bestimmung von Interkorrelationskoeffizienten zwischen den beiden synthetisierten Szenen an bestimmten Umgebungen der Punkte der Szene verglichen werden, und daß die erhaltenen Koeffizienten mit einem Schwellenwert verglichen werden.
3. Verfahren nach Anspruch 1 oder 2, dadurch gekennzeichnet, daß das geometrische Modell der Szene ausgehend von in einem Bezugszustand von den Sensoren gelieferten, pixelweise gespeicherten Abbildungen der Szene erstellt wird.
4. Verfahren nach Anspruch 1 oder 2, dadurch gekennzeichnet, daß das geometrische Modell in mathematischer Form ausgehend von einer durch vorhergegangene Beobachtungen oder mit Hilfe von Karten zur Verfügung gestellten Kenntnis des Geländes gebildet wird.
5. Verfahren nach Anspruch 1 oder 2, dadurch gekennzeichnet, daß das geometrische Modell der Szene ausgehend von Landmarkenpunkten mit bekannten Koordinaten (x,y,z) und einer Gleichung mit einer vorgegebenen Form erstellt wird.
6. Verfahren nach Anspruch 1 oder 2, dadurch gekennzeichnet, daß das geometrische Modell ausgehend von einem Gitternetz erstellt wird, welches durch Abtasten mit Hilfe eines Lasers über die Szene gelegt und in Form der Ortung des bei jedem Laserimpuls den Auftreffpunkt auf dem Gelände darstellenden Bildpunktes durch jeden der Sensoren registriert wird.
7. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß des weiteren aufeinanderfolgende, von den Sensoren gelieferte Abbildungen zum Erstellen einer zeitlichen Erfassung verglichen werden.
8. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß die Sensoren einen gegenseitigen Abstand haben, der geringer als die Entfernung zwischen den Sensoren und den nächstliegenden Punkten der Szene ist.
9. Vorrichtung zur Überwachung einer angestrahlten dreidimensionalen Szene mit einer bekannten oder bestimmbaren geometrischen Struktur, welche die Erfassung von Reliefanomalien auf der Szene ermöglicht, mit mindestens zwei Sensoren (10, 12), welche in Bezug auf einen radiometrischen Parameter in vergleichbaren Wellenlängen empfindlich sind, und des weiteren mit: Einrichtungen zum Speichern eines rein geometrischen dreidimensionalen Bezugsmodells der zu überwachenden Szene, welche jedem Punkt der Szene homologe Punkte im Feld der Sensoren zuordnen; Einrichtungen, welche das Synthetisieren der Szenen ermöglichen, zur Herstellung einer Anzahl von synthetisierten Szenen, die gleich derjenigen der Sensoren ist, wobei jede der Szenen ausgehend von dem geometrischen Modell sowie Signalen, welche repräsentativ für den radiometrischen Wert jedes der von

EP 0 577 491 B1

einem der Sensoren gelieferten Bildpunkte sind, durch Projizieren auf das dreidimensionale geometrische Modell synthetisiert wird; und Einrichtungen zum Vergleichen von Umgebungen mit vorgegebener Abmessung auf der Gesamtheit der synthetisierten Szenen ausgehend von den von den verschiedenen Sensoren gelieferten Signalen, wodurch die Bestimmung der Abweichungen zwischen den gleichen Umgebungen in den beiden synthetisierten Szenen ermöglicht wird.

10. Vorrichtung nach Anspruch 9, dadurch gekennzeichnet, daß die Sensoren des weiteren Sensoren aufweisen, welche in Wellenlängenbereichen arbeiten, welche von denjenigen der beiden Sensoren verschieden sind.
11. Überwachungs Vorrichtung nach Anspruch 10 zum Lokalisieren eines mobilen Robots in der Szene, dadurch gekennzeichnet, daß die Sensoren von Einrichtungen zum Verfahren in Bezug auf die Szene getragen sind, und mit einem Speicher zum Aktualisieren des gespeicherten geometrischen Modells in Abhängigkeit von der Position der Sensoren in Bezug auf die Szene verbunden sind.

Claims

1. A method for the surveillance of a three-dimensional scene having a known or determinable geometrical structure, making it possible to detect the appearance of an anomaly in relief and including the following stages:
 - a three-dimensional geometrical reference model of the said scene is made up and memorised, in a reference state, associating with each point of the scene homologous points in the field of at least two sensors placed at a distance from each other and sensitive to a radiometric parameter (such as the brilliance) of the scene;
 - periodically, the real scene is synthesised over the three-dimensional geometrical model on a basis of each of the representations thereof which are furnished by the sensors in order to obtain as many synthesised scenes as there are sensors, and
 - points of closeness belonging to the synthesised scenes are compared with one another in order to show up divergencies due to a modification of the three-dimensional geometrical model, only those divergencies which correspond to a minimal height anomaly in excess of a specific value being retained.
2. A method according to claim 1, characterised in that the synthesised scenes are compared by determining coefficients of intercorrelation between the two synthesised scenes over specific closenesses of the points of the scene and in that the coefficients obtained are compared with a threshold.
3. A method according to claim 1 or 2, characterised in that the geometrical model of the scene is made up on the basis of images of the scene, in a reference state provided by the sensors and memorised pixel by pixel.
4. A method according to claim 1 or 2, characterised in that the geometrical model is established in mathematical form on the basis of knowledge of the terrain supplied by prior observations or with the help of maps.
5. A method according to claim 1 or 2, characterised in that the geometrical model of the scene is made up on the basis of landmark points of known co-ordinates (x, y, z) and an equation of predetermined form
6. A method according to claim 1 or 2, characterised in that the geometrical model is made up on the basis of a currying grid produced on the scene by scanning using a laser and recorded in the form of the location for each sensor, at each laser pulse, in respect of that point of the image which represents the impact on the terrain.
7. A method according to any one of the preceding claims, characterised in that, in addition, successive images delivered by the sensors are compared in order to carry out a time-related detection.
8. A method according to any one of the preceding claims, characterised in that the sensors are spaced apart mutually by an amount less than the distance between the sensors and the closest points of the scene.
9. A device for surveillance of a three-dimensional illuminated scene having a known or determinable geometrical structure, making it possible to detect relief anomalies over the scene, comprising at least two sensors (10, 12) sensitive to a radiometric parameter, in comparable wavelengths, and comprising likewise: means of memorising a three-dimensional purely geometrical reference model of the scene to be surveyed, associating with each point of the scene homologous points in the field of the sensors; means which make it possible to synthesise scenes in

order to obtain as many synthesised scenes as there are sensors, each of the scenes being synthesised on the basis of the geometrical model and signals representing the radiometric value of each image pixel supplied by one of the sensors by projection on the three-dimensional geometrical model, and means of comparing points of closeness of predetermined size on the set of synthetic scenes on the basis of signals delivered by the various sensors, making it possible to determine the distances between the same locations in the two synthesised scenes.

10. A device according to claim 9, characterised in that the said sensors furthermore comprise sensors working in different wavelength ranges over the two said sensors.

11. A surveillance device according to claim 10 intended to locate a robot adapted for movement within the scene, characterised in that the said sensors are carried by means of displacement in respect of the scene and are associated with a memory for updating the geometrical model memorised as a function of the position of the sensors in respect of the scene.

FIG.1.

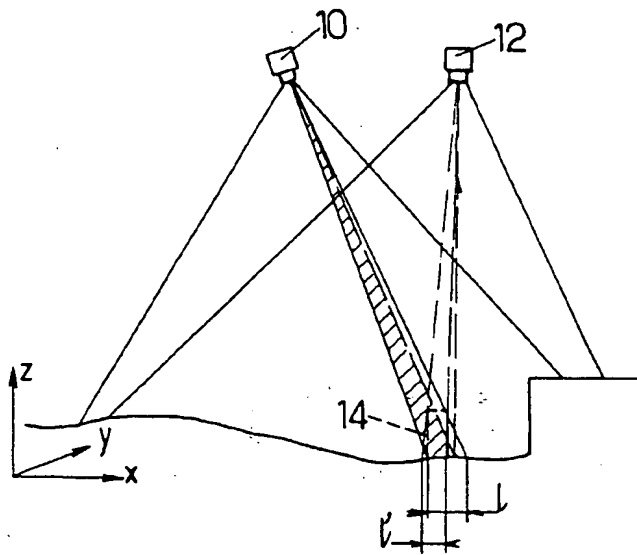


FIG.2.

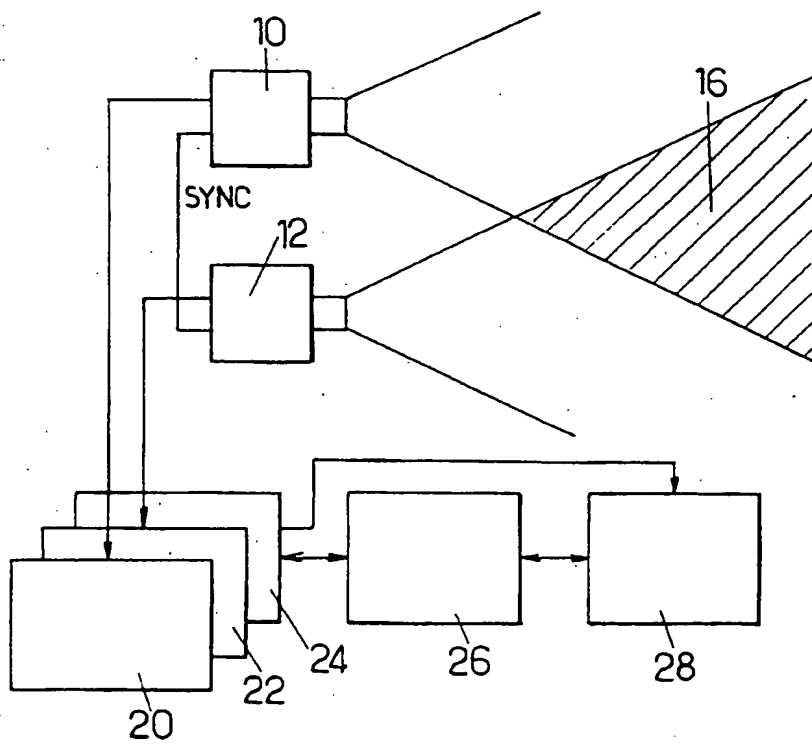
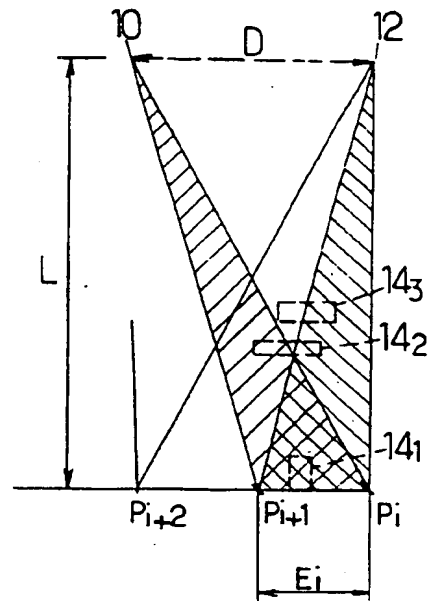


FIG.4.

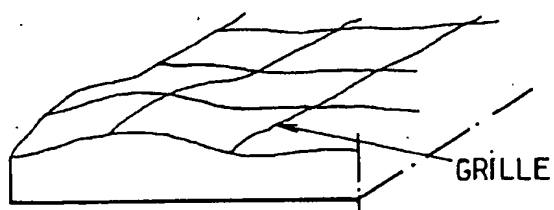
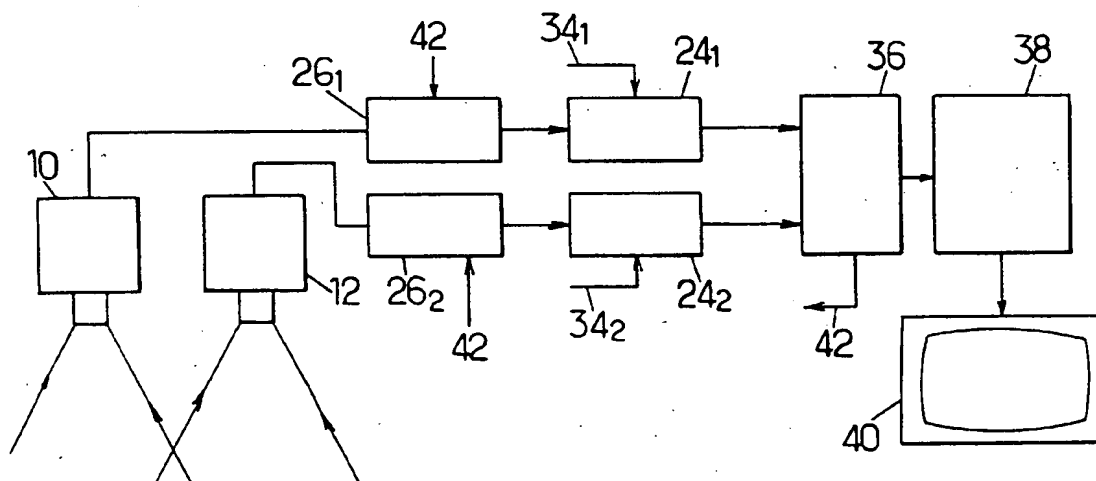
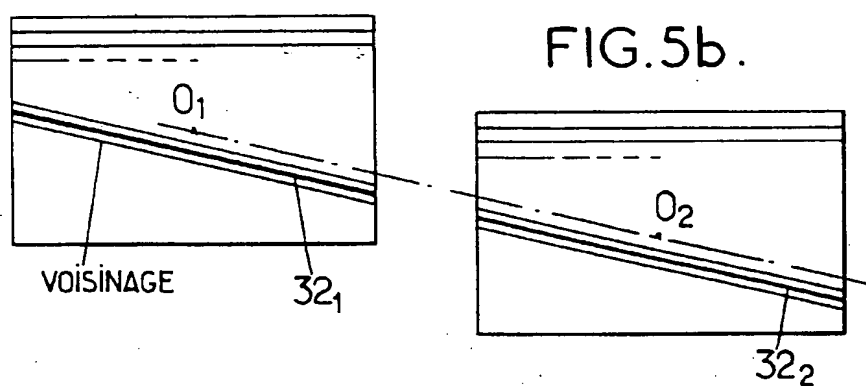
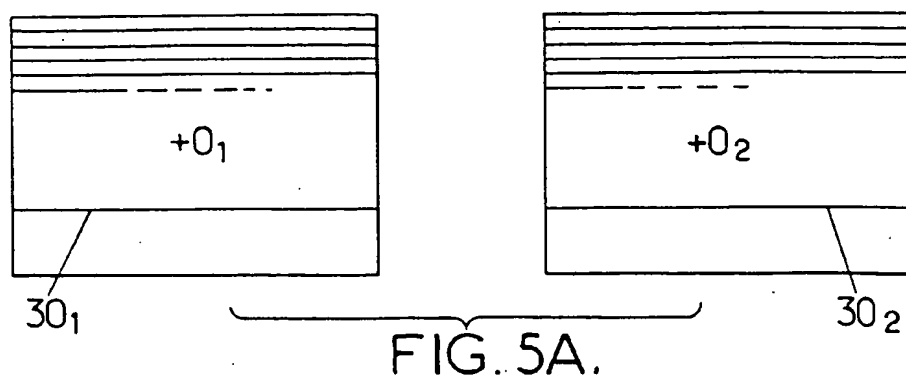


FIG.3.

